RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/656,309**

DATE: 04/16/2001
TIME: 11:25:44

Input Set : **A:\es.txt**
Output Set: **N:\CRF3\04162001\I656309.raw**

ENTERED

```
 3 <110> APPLICANT: DIVERSA CORPORATION
 4        CALLEN, Walter
 5        MATHUR, Eric
 7 <120> TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND
METHODS OF USE
 8        THEREOF
10 <130> FILE REFERENCE: DIVER1350-2
12 <140> CURRENT APPLICATION NUMBER: US 09/656,309
13 <141> CURRENT FILING DATE: 2000-09-06
15 <160> NUMBER OF SEQ ID NOS: 2
17 <170> SOFTWARE: PatentIn version 3.0
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 2412
21 <212> TYPE: DNA
22 <213> ORGANISM: Pyrolobus fumarius
24 <220> FEATURE:
25 <221> NAME/KEY: CDS
26 <222> LOCATION: (1)..(2412)
28 <400> SEQUENCE: 1
29 atg act gaa gtt gta ttc acg gtt tta gac tct agc tac gag gtt gtt       48
30 Met Thr Glu Val Val Phe Thr Val Leu Asp Ser Ser Tyr Glu Val Val
31 1               5                   10                  15
33 ggt aaa gag cct cag gta atc ata tgg ggt att gct gag aac ggc gag       96
34 Gly Lys Glu Pro Gln Val Ile Ile Trp Gly Ile Ala Glu Asn Gly Glu
35               20                  25                  30
37 agg gta gtc ctc att gac agg tct ttt cgc cca tac ttc tat gcg ctg      144
38 Arg Val Val Leu Ile Asp Arg Ser Phe Arg Pro Tyr Phe Tyr Ala Leu
39           35                  40                  45
41 ctt gca ccg ggc gcc gat cct aag cag gta gca caa cgt att cgt gca      192
42 Leu Ala Pro Gly Ala Asp Pro Lys Gln Val Ala Gln Arg Ile Arg Ala
43       50                  55                  60
45 ttg agt agg cca aag agc ccg att ata ggt gta gag gat gac aag agg      240
46 Leu Ser Arg Pro Lys Ser Pro Ile Ile Gly Val Glu Asp Asp Lys Arg
47 65                  70                  75                  80
49 aag tac ttc ggg agg cct cgt agg gtc tta cgt att cgc acc gtg cta      288
50 Lys Tyr Phe Gly Arg Pro Arg Arg Val Leu Arg Ile Arg Thr Val Leu
51               85                  90                  95
53 ccc gag gct gtt agg gag tat cgc gaa ctc gta aag aac gtt gat ggt      336
54 Pro Glu Ala Val Arg Glu Tyr Arg Glu Leu Val Lys Asn Val Asp Gly
55               100                 105                 110
57 gtt gag gat gtt cta gag gcg gat ata cgc ttc gct atg cgc tat ctc      384
58 Val Glu Asp Val Leu Glu Ala Asp Ile Arg Phe Ala Met Arg Tyr Leu
59           115                 120                 125
61 ata gat cac gat cta ttt cct ttc acc tgg tac cgt gta gag gct gag      432
62 Ile Asp His Asp Leu Phe Pro Phe Thr Trp Tyr Arg Val Glu Ala Glu
63       130                 135                 140
65 ccc ctc gag aac aag atg ggc ttc cgt gtc gac aag gta tac ctg gtt      480
66 Pro Leu Glu Asn Lys Met Gly Phe Arg Val Asp Lys Val Tyr Leu Val
```

RAW SEQUENCE LISTING   DATE: 04/16/2001
PATENT APPLICATION: US/09/656,309   TIME: 11:25:44

Input Set : A:\es.txt
Output Set: N:\CRF3\04162001\I656309.raw

```
67 145                 150                 155                 160
69 aag agc agg ccg gag cca ctt tat ggt gag gct ctc gca cca acc aag    528
70 Lys Ser Arg Pro Glu Pro Leu Tyr Gly Glu Ala Leu Ala Pro Thr Lys
71             165                 170                 175
73 ctt ccc gat ctt agg ata ctc gcg ttc gat att gaa gtt tat agc aag    576
74 Leu Pro Asp Leu Arg Ile Leu Ala Phe Asp Ile Glu Val Tyr Ser Lys
75             180                 185                 190
77 caa ggg tcg ccg cgt cca gag cgc gat cct gta ata gtg ata gct gtg    624
78 Gln Gly Ser Pro Arg Pro Glu Arg Asp Pro Val Ile Val Ile Ala Val
79             195                 200                 205
81 aag act gac gat ggc gat gag gtg cta ttc att gca gag ggc aaa gac    672
82 Lys Thr Asp Asp Gly Asp Glu Val Leu Phe Ile Ala Glu Gly Lys Asp
83       210                 215                 220
85 gat cga aaa ccg ata cgc gag ttt gta gag tac gtg aag agg tat gac    720
86 Asp Arg Lys Pro Ile Arg Glu Phe Val Glu Tyr Val Lys Arg Tyr Asp
87 225                 230                 235                 240
89 ccc gac ata ata gtc ggt tat aac aac aat cat ttc gat tgg cct tat    768
90 Pro Asp Ile Ile Val Gly Tyr Asn Asn Asn His Phe Asp Trp Pro Tyr
91             245                 250                 255
93 ctt ttg agg cgc gcc cgc atc cta ggc ata aag ctt gat gtg act aga    816
94 Leu Leu Arg Arg Ala Arg Ile Leu Gly Ile Lys Leu Asp Val Thr Arg
95             260                 265                 270
97 aga gtt ggc gcc gag ccc acc act agc gta cat ggg cac gtc tct gtc    864
98 Arg Val Gly Ala Glu Pro Thr Thr Ser Val His Gly His Val Ser Val
99             275                 280                 285
101 cct ggc agg ctt aac gta gat ctg tac gac tat gcc gaa gag atg cca    912
102 Pro Gly Arg Leu Asn Val Asp Leu Tyr Asp Tyr Ala Glu Glu Met Pro
103      290                 295                 300
105 gag atc aag ata aag agt ctc gag gag gtc gca gag tat cta ggc gtg    960
106 Glu Ile Lys Ile Lys Ser Leu Glu Glu Val Ala Glu Tyr Leu Gly Val
107 305                 310                 315                 320
109 atg aag aag agt gaa cgc gtt atc atc aat tgg tgg gag att cca gac    1008
110 Met Lys Lys Ser Glu Arg Val Ile Ile Asn Trp Trp Glu Ile Pro Asp
111             325                 330                 335
113 tat tgg gac gac ccg aag aag aga cca cta tta ctg caa tac gcg cgc    1056
114 Tyr Trp Asp Asp Pro Lys Lys Arg Pro Leu Leu Leu Gln Tyr Ala Arg
115             340                 345                 350
117 gac gat gtc cgc gct act tac ggc tta gcc gag aag ata ttg ccg ttt    1104
118 Asp Asp Val Arg Ala Thr Tyr Gly Leu Ala Glu Lys Ile Leu Pro Phe
119             355                 360                 365
121 gct atc cag ttg tcg tac gta aca ggt ctc cca cta gac cag gta ggt    1152
122 Ala Ile Gln Leu Ser Tyr Val Thr Gly Leu Pro Leu Asp Gln Val Gly
123      370                 375                 380
125 gcg atg agt gtt ggc ttt cga ctt gaa tgg tac ctg ata cgc gcg gcg    1200
126 Ala Met Ser Val Gly Phe Arg Leu Glu Trp Tyr Leu Ile Arg Ala Ala
127 385                 390                 395                 400
129 ttt aag atg aaa gag ctt gtg ccg aac cgc gtt gag cgc cca gaa gag    1248
130 Phe Lys Met Lys Glu Leu Val Pro Asn Arg Val Glu Arg Pro Glu Glu
131             405                 410                 415
```

RAW SEQUENCE LISTING   DATE: 04/16/2001
PATENT APPLICATION: US/09/656,309   TIME: 11:25:44

Input Set : A:\es.txt
Output Set: N:\CRF3\04162001\I656309.raw

```
133 act tac cgt ggc gct ata gtt ctt gag ccg ttg aga ggc gtg cac gag     1296
134 Thr Tyr Arg Gly Ala Ile Val Leu Glu Pro Leu Arg Gly Val His Glu
135         420                 425                 430
137 aat ata gcc gta ctc gac ttt agc tcg atg tac cca aac atc atg ata     1344
138 Asn Ile Ala Val Leu Asp Phe Ser Ser Met Tyr Pro Asn Ile Met Ile
139         435                 440                 445
141 aag tac aat gtt ggt cct gac acg ctt gtg agg cct ggt gaa aag tgt     1392
142 Lys Tyr Asn Val Gly Pro Asp Thr Leu Val Arg Pro Gly Glu Lys Cys
143     450                 455                 460
145 ggc gag tgt ggt tgc tgg gag gcc ccg gag gtc aag cac agg ttc cgt     1440
146 Gly Glu Cys Gly Cys Trp Glu Ala Pro Glu Val Lys His Arg Phe Arg
147 465                 470                 475                 480
149 agg tgt ccg ccc ggc ttc ttc aag aca gtt ctt gag agg ctg tta gag     1488
150 Arg Cys Pro Pro Gly Phe Phe Lys Thr Val Leu Glu Arg Leu Leu Glu
151         485                 490                 495
153 ctt cgt aag cgt gtg cgt gct gaa atg aag aag tat cct ccg gat agc     1536
154 Leu Arg Lys Arg Val Arg Ala Glu Met Lys Lys Tyr Pro Pro Asp Ser
155         500                 505                 510
157 cca gaa tat cga ctg ttg gat gaa agg cag aag gcg ttg aag gtt ctt     1584
158 Pro Glu Tyr Arg Leu Leu Asp Glu Arg Gln Lys Ala Leu Lys Val Leu
159         515                 520                 525
161 gca aac gct agt tac ggc tac atg ggt tgg agc ggc gct agg tgg tat     1632
162 Ala Asn Ala Ser Tyr Gly Tyr Met Gly Trp Ser Gly Ala Arg Trp Tyr
163     530                 535                 540
165 tgc agg gag tgc gca aag gct gtc acg gct tgg ggt agg cac ctc ata     1680
166 Cys Arg Glu Cys Ala Lys Ala Val Thr Ala Trp Gly Arg His Leu Ile
167 545                 550                 555                 560
169 cgc acc gcc atc aac ata gct cgt aaa cta ggc ctc aag gtg atc tac     1728
170 Arg Thr Ala Ile Asn Ile Ala Arg Lys Leu Gly Leu Lys Val Ile Tyr
171         565                 570                 575
173 ggt gac aca gat tcg ctc ttc gtg acc tat gat ccg gag aag gtg gaa     1776
174 Gly Asp Thr Asp Ser Leu Phe Val Thr Tyr Asp Pro Glu Lys Val Glu
175         580                 585                 590
177 aat ttc atc aaa att ata aag gag gag ctg ggg ttc gaa atc aag cta     1824
178 Asn Phe Ile Lys Ile Ile Lys Glu Glu Leu Gly Phe Glu Ile Lys Leu
179         595                 600                 605
181 gag aag gtg tac aaa cgc tta ttc ttt aca gag gct aag aag agg tac     1872
182 Glu Lys Val Tyr Lys Arg Leu Phe Phe Thr Glu Ala Lys Lys Arg Tyr
183     610                 615                 620
185 gct ggc ctt ctc gag gac gga cgt ata gat att gtc ggt ttc gag gct     1920
186 Ala Gly Leu Leu Glu Asp Gly Arg Ile Asp Ile Val Gly Phe Glu Ala
187 625                 630                 635                 640
189 gta cgt ggc gat tgg tgt gaa ctc gcc aag gag gtt cag act aag gtt     1968
190 Val Arg Gly Asp Trp Cys Glu Leu Ala Lys Glu Val Gln Thr Lys Val
191                 645                 650                 655
193 gtc gaa ata gta ttg aag acg agt gag gtg aac aag gct gta gag tac     2016
194 Val Glu Ile Val Leu Lys Thr Ser Glu Val Asn Lys Ala Val Glu Tyr
195         660                 665                 670
197 gtc agg aag att gtg aaa gag ttg gag gag ggc aag gtt ccc ata gag     2064
```

```
                RAW SEQUENCE LISTING            DATE: 04/16/2001
                PATENT APPLICATION:  US/09/656,309    TIME: 11:25:44

                Input Set : A:\es.txt
                Output Set: N:\CRF3\04162001\I656309.raw

198 Val Arg Lys Ile Val Lys Glu Leu Glu Glu Gly Lys Val Pro Ile Glu
199         675                 680                 685
201 aag ctt gta atc tgg aag acc ctt agt aag cgt ctt gag gag tac aca    2112
202 Lys Leu Val Ile Trp Lys Thr Leu Ser Lys Arg Leu Glu Glu Tyr Thr
203         690                 695                 700
205 acg gag gca cca cac gtc gtt gca gcg aag agg atg ctg tca gca ggc    2160
206 Thr Glu Ala Pro His Val Val Ala Ala Lys Arg Met Leu Ser Ala Gly
207 705                 710                 715                 720
209 tac cgg gta agc cca ggc gac aag ata ggg tat gta ata gtg aag ggt    2208
210 Tyr Arg Val Ser Pro Gly Asp Lys Ile Gly Tyr Val Ile Val Lys Gly
211                 725                 730                 735
213 ggt ggc cgt atc agt caa aga gca tgg cca tac ttc atg gtc aag gat    2256
214 Gly Gly Arg Ile Ser Gln Arg Ala Trp Pro Tyr Phe Met Val Lys Asp
215                 740                 745                 750
217 cct agc cag ata gac gtg acc tac tat gtt gac cac caa atc atc ccg    2304
218 Pro Ser Gln Ile Asp Val Thr Tyr Tyr Val Asp His Gln Ile Ile Pro
219                 755                 760                 765
221 gct gca ttg aga ata ctg ggc tac ttt ggc atc acc gag aag aag ctg    2352
222 Ala Ala Leu Arg Ile Leu Gly Tyr Phe Gly Ile Thr Glu Lys Lys Leu
223         770                 .775                780
225 aaa gca agt gca act ggg cag aag act ctc ttc gac ttt cta gcc aag    2400
226 Lys Ala Ser Ala Thr Gly Gln Lys Thr Leu Phe Asp Phe Leu Ala Lys
227 785                 790                 795                 800
229 aag agc aag taa                                                    2412
230 Lys Ser Lys
234 <210> SEQ ID NO: 2
235 <211> LENGTH: 803
236 <212> TYPE: PRT
237 <213> ORGANISM: Pyrolobus fumarius
239 <400> SEQUENCE: 2
241 Met Thr Glu Val Val Phe Thr Val Leu Asp Ser Ser Tyr Glu Val Val
242 1               5                   10                  15
245 Gly Lys Glu Pro Gln Val Ile Ile Trp Gly Ile Ala Glu Asn Gly Glu
246             20                  25                  30
249 Arg Val Val Leu Ile Asp Arg Ser Phe Arg Pro Tyr Phe Tyr Ala Leu
250         35                  40                  45
253 Leu Ala Pro Gly Ala Asp Pro Lys Gln Val Ala Gln Arg Ile Arg Ala
254     50                  55                  60
257 Leu Ser Arg Pro Lys Ser Pro Ile Ile Gly Val Glu Asp Asp Lys Arg
258 65                  70                  75                  80
261 Lys Tyr Phe Gly Arg Pro Arg Arg Val Leu Arg Ile Arg Thr Val Leu
262             85                  90                  95
265 Pro Glu Ala Val Arg Glu Tyr Arg Glu Leu Val Lys Asn Val Asp Gly
266             100                 105                 110
269 Val Glu Asp Val Leu Glu Ala Asp Ile Arg Phe Ala Met Arg Tyr Leu
270         115                 120                 125
273 Ile Asp His Asp Leu Phe Pro Phe Thr Trp Tyr Arg Val Glu Ala Glu
274         130                 135                 140
277 Pro Leu Glu Asn Lys Met Gly Phe Arg Val Asp Lys Val Tyr Leu Val
```

RAW SEQUENCE LISTING   DATE: 04/16/2001
PATENT APPLICATION: US/09/656,309   TIME: 11:25:44

Input Set : A:\es.txt
Output Set: N:\CRF3\04162001\I656309.raw

```
278 145                 150                 155                 160
281 Lys Ser Arg Pro Glu Pro Leu Tyr Gly Glu Ala Leu Ala Pro Thr Lys
282             165                 170                 175
285 Leu Pro Asp Leu Arg Ile Leu Ala Phe Asp Ile Glu Val Tyr Ser Lys
286             180                 185                 190
289 Gln Gly Ser Pro Arg Pro Glu Arg Asp Pro Val Ile Val Ile Ala Val
290             195                 200                 205
293 Lys Thr Asp Asp Gly Asp Glu Val Leu Phe Ile Ala Glu Gly Lys Asp
294     210                 215                 220
297 Asp Arg Lys Pro Ile Arg Glu Phe Val Glu Tyr Val Lys Arg Tyr Asp
298 225                 230                 235                 240
301 Pro Asp Ile Ile Val Gly Tyr Asn Asn Asn His Phe Asp Trp Pro Tyr
302             245                 250                 255
305 Leu Leu Arg Arg Ala Arg Ile Leu Gly Ile Lys Leu Asp Val Thr Arg
306             260                 265                 270
309 Arg Val Gly Ala Glu Pro Thr Thr Ser Val His Gly His Val Ser Val
310             275                 280                 285
313 Pro Gly Arg Leu Asn Val Asp Leu Tyr Asp Tyr Ala Glu Glu Met Pro
314     290                 295                 300
317 Glu Ile Lys Ile Lys Ser Leu Glu Glu Val Ala Glu Tyr Leu Gly Val
318 305                 310                 315                 320
321 Met Lys Lys Ser Glu Arg Val Ile Ile Asn Trp Trp Glu Ile Pro Asp
322             325                 330                 335
325 Tyr Trp Asp Asp Pro Lys Lys Arg Pro Leu Leu Leu Gln Tyr Ala Arg
326             340                 345                 350
329 Asp Asp Val Arg Ala Thr Tyr Gly Leu Ala Glu Lys Ile Leu Pro Phe
330             355                 360                 365
333 Ala Ile Gln Leu Ser Tyr Val Thr Gly Leu Pro Leu Asp Gln Val Gly
334     370                 375                 380
337 Ala Met Ser Val Gly Phe Arg Leu Glu Trp Tyr Leu Ile Arg Ala Ala
338 385                 390                 395                 400
341 Phe Lys Met Lys Glu Leu Val Pro Asn Arg Val Glu Arg Pro Glu Glu
342             405                 410                 415
345 Thr Tyr Arg Gly Ala Ile Val Leu Glu Pro Leu Arg Gly Val His Glu
346             420                 425                 430
349 Asn Ile Ala Val Leu Asp Phe Ser Ser Met Tyr Pro Asn Ile Met Ile
350             435                 440                 445
353 Lys Tyr Asn Val Gly Pro Asp Thr Leu Val Arg Pro Gly Glu Lys Cys
354     450                 455                 460
357 Gly Glu Cys Gly Cys Trp Glu Ala Pro Glu Val Lys His Arg Phe Arg
358 465                 470                 475                 480
361 Arg Cys Pro Pro Gly Phe Phe Lys Thr Val Leu Glu Arg Leu Leu Glu
362             485                 490                 495
365 Leu Arg Lys Arg Val Arg Ala Glu Met Lys Lys Tyr Pro Pro Asp Ser
366             500                 505                 510
369 Pro Glu Tyr Arg Leu Leu Asp Glu Arg Gln Lys Ala Leu Lys Val Leu
370             515                 520                 525
373 Ala Asn Ala Ser Tyr Gly Tyr Met Gly Trp Ser Gly Ala Arg Trp Tyr
374     530                 535                 540
```

VERIFICATION SUMMARY   DATE: 04/16/2001
PATENT APPLICATION: **US/09/656,309**   TIME: 11:25:45

Input Set : **A:\es.txt**
Output Set: **N:\CRF3\04162001\I656309.raw**